
CHAPTER EIGHT: ANNOTATED BIBLIOGRAPHY

In selecting publications for inclusion in this short bibliography, an effort was made to incorporate those which can be useful for evaluators who want to find information relevant to the tasks they will be faced with and which this brief handbook could not cover in depth. Thus, we have not necessarily included all books which might be on a reading list in a graduate course dealing with educational evaluation, but have selected those which NSF/HRD grantees should find most useful. This includes most of those referenced in Chapters 1-4, as well as others which provide perspectives and methods not extensively covered in the handbook.

We have divided this bibliography into two sections. Section One includes books that deal in large part with the broad topics of theory and practice in the field of evaluation, with emphasis on those which deal specifically with education. Section Two is more narrowly focused on technical topics and/or single issues. Of course, this is not an airtight division: many of the broad-based works contain a great deal of technical information and hands-on advice.

Section One: Theory and Practice

House, Ernest R. (1993). *Professional Evaluation — Social Impact and Political Consequences*. Newbury Park, CA: Sage.

The author is a professor in the School of Education at the University of Colorado and an experienced evaluator of major social and educational programs. In the author's own words, this book: "is about analyzing the social, political, economic, historical, and cultural influences on evaluation." It provides a thoughtful overview of the field's evolution, from its earlier reliance on "value-free" experimental and primarily quantitative methods to the current emphasis on methods more appropriate to the diverse, multicultural and politically charged environment in which social and educational programs operate.

Rossi, Peter H. & Freeman, Howard E. (1993). *Evaluation — A Systematic Approach* (5th Edition). Newbury Park, CA: Sage.

This is the newest edition of one of the most comprehensive and widely used texts about evaluation. It provides extensive and sophisticated discussions of all aspects of designing and assessing the implementation and utility of social programs. Many of the projects cited and discussed in this volume deal with educational programs and innovations, although the bulk of the programs with which these authors had first hand experience were in the fields of housing, health services, and criminal justice. Most tasks that evaluators are likely to be asked to perform, and most problems they will have to deal with, technical as well as political, are covered here. The authors adhere to the social science model in their approach to evaluation, with clear preference for randomized and quasi-experimental designs, but they also cover other evaluation methods, including the use of qualitative and judgmental approaches.

Worthen, Blaine R. & Sanders, James R. (1987). *Educational Evaluation: Alternative Approaches and Practical Guidelines*. White Plains, NY: Longman.

This book was designed primarily as a basic text for graduate courses in evaluation, or related administration, curriculum, or teacher education courses, where efforts are made to teach practitioners how to assess the effectiveness of their educational endeavors. It seeks to familiarize readers with alternative approaches for planning evaluations, and provides step-by-step practical guidelines for conducting them. The book is very systematically organized, describes at length all evaluation approaches which have been developed over the years, and includes a great deal of useful information, especially for the inexperienced evaluator. A detailed, naturalistic description of the conduct of an evaluation program, including problems encountered with school staff, other stakeholders, and administrators, provides a useful example of “real world” issues in evaluation.

Scriven, Michael (1991). *Evaluation Thesaurus (4th Edition)*. Newbury Park, CA: Sage.

A highly original, wide ranging collection of ideas, concepts, positions and techniques which reflect the critical, incisive, and often unconventional views held by this leader in the field of evaluation. It contains a 40-page introductory essay on the nature of evaluation and nearly 1,000 entries which range from one paragraph definitions of technical terms and acronyms to philosophical and methodological discussions extending over many pages. The Thesaurus is not focused on the field of education, but it provides excellent coverage of issues and concepts of interest to educational evaluators.

Guba, E. G. & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage.

The authors propose a monumental shift in evaluation practice, advocating the constructionist position in its most extreme form. Guba and Lincoln describe problems faced by previous generations of evaluators—politics, ethical dilemmas, imperfections and gaps, inclusive deductions—and lay the blame for failure and nonutilization at the feet of the unquestioned reliance on the scientific/positivist paradigm of research. Fourth generation evaluation moves beyond science to include the myriad human, political, social, cultural, and contextual elements that are involved in the evaluation process.

The book describes the differences between the earlier (and still widely used) evaluation model, based on positivist/scientific assumptions and statistical techniques, and the “naturalistic” approach to evaluation, and outlines methodological guidelines for the conduct of naturalistic evaluations.

Joint Committee on Standards for Educational Evaluation (1981). *Standards for Evaluations of Educational Programs, Projects, and Materials*. New York, NY: McGraw Hill.

For more than 10 years this volume has provided guidance to educational evaluators. It contains 30 standards for program evaluation along with examples of their application. It is the authoritative reference on principles of good and ethical program evaluation. Sponsored by several professional associations concerned with the quality of program evaluations in education, the Joint Committee has defined state-of-the-art principles, guidelines, and illustrative cases that can be used to judge the quality of any evaluation. The standards fall into four categories: utility, feasibility, propriety, and accuracy.

Patton, Michael Q. (1986). *Utilization-Focused Evaluation (2nd Edition)*. Newbury Park, CA: Sage.

In a book that combines both the theoretical and the practical, Patton examines how and why to conduct evaluations. In this revised and updated edition, the author provides practical advice grounded in evaluation theory and practice, and shows how to conduct evaluations from beginning to end in ways that will be useful—and used. This volume discusses the ferment and changes in evaluation during the eighties and the tremendous growth of “in-house” evaluations conducted by internal evaluators. Patton also discusses a methodological synthesis of the “qualitative versus the quantitative” methods debate, as well as the cross-cultural development of evaluation as an internationally recognized profession. Presented in an integrated framework, the author discusses topics such as utilization built on a new, concise definition of evaluation which emphasizes providing useful and usable information to specific people.

Section Two: Technical Topics

Jaeger, R. M. (1990). *Statistics: A Spectator Sport (2nd Edition)*. Newbury Park, CA: Sage.

This book takes the reader to the point of understanding advanced statistics without introducing complex formulas or equations. It covers most of the statistical concepts and techniques which evaluators commonly use in the design and analysis of evaluation studies, and most of the examples and illustrations are from actual studies performed in the field of education. The topics included range from descriptive statistics, including measures of central tendency and fundamentals of measurement, to inferential statistics and advanced analytic methods.

Campbell, Donald T. & Stanley, Julian C. (1966). *Experimental and Quasi-experimental Designs for Research*. Boston, MA: Houghton Mifflin.

This slim (84 pages) volume is a slightly enlarged version of the chapter originally published in the 1963 *Handbook of Research on Teaching* and is considered the classic text on valid experimental and quasi-experimental designs in real world situations where the experimenter has very limited control over the environment. To this day, it is the most useful basic reference book for evaluators who plan the use of such designs.

Herman, Joan L. (Editor, 1987). *Program Evaluation Kit*. Newbury Park, CA: Sage.

This kit, prepared by the Center for the Study of Evaluation at the University of California, Los Angeles, contains nine books written to guide and assist evaluators in planning and executing evaluations, with emphasis on practical, field-tested step-by-step procedures and with considerable attention to the management of each phase. The kit makes heavy use of charts, illustrations, and examples to clarify the material for novice evaluators. Volume 1, ***Evaluator's Handbook***, provides an overview of evaluation activities, describes the evaluation perspective which guides the kit, and discusses specific procedures for conducting Formative and Summative evaluations. The remaining eight volumes deal with specific topics:

Volume Two ***How to Focus an Evaluation***

Volume Three ***How to Design a Program Evaluation***

Volume Four ***How to Use Quantitative Methods in Evaluation***

Volume Five ***How to Assess Program Implementation***

Volume Six ***How to Measure Attitudes***

Volume Seven ***How to Measure Performance and Use Tests***

Volume Eight ***Hot to Analyze Data***

Volume Nine ***How to Communicate Evaluation Findings***

Depending on their needs, evaluators will find every one of these volumes useful. Volume Seven, ***How to Measure Performance and Use Tests***, covers a topic for which we have not located another suitable text for inclusion in this bibliography.

The Kit can be purchased as a unit, or by ordering individual volumes.

Yin, Robert K. (1989). *Case Study Research: Design and Method*. Newbury Park, CA: Sage.

The author's background in experimental psychology may explain the emphasis in this book on the use of rigorous methods in the conduct and analysis of case studies, thus minimizing what many believe is a spurious distinction between quantitative and qualitative studies. While arguing eloquently that case studies are an important tool when an investigator (or evaluator) has little control over events and when the focus is on a contemporary phenomenon within some real-life context, the author insists that case studies be designed and analyzed so as to provide generalizable findings. Although the focus is on design and analysis, data collection and report writing are also covered.

Fowler, Floyd J., Jr. (1993). *Survey Research Methods* (2nd Edition). Newbury Park, CA: Sage.

Using non-technical language, the author has provided a comprehensive discussion of survey design (including sampling, data collection methods, and the design of survey questions) and procedures which constitute good survey practice, including attention to data quality and ethical issues. According to the author, "this book is intended to provide perspective and understanding to those who would be designers or users of survey research, at the same time as it provides a sound first step for those who actually may go about collecting data."

Stewart, David W. & Shamdasani, Prem N. (1990). *Focus Groups: Theory and Practice*. Newbury Park, CA: Sage.

This book differs from many others published in recent years which address primarily techniques of recruiting participants and the actual conduct of focus group sessions. This volume pays considerable attention to the fact that focus groups are by definition an exercise in group dynamics, which must be taken into account when conducting such groups and especially when interpreting the results obtained. However, it also covers very adequately practical issues such as recruitment of participants, the role of the moderator, and appropriate techniques for data analysis.

Linn, Robert L., Baker, Eva L., & Dunbar, Stephen B. (1991). Complex, performance-based assessment: expectations and validation criteria. *Educational Researcher*, (Vol. 20, No. 8, pp. 15-22).

In recent years there has been an increasing emphasis on assessment results, as well as increasing concern about the nature of the most widely used forms of student assessment and uses that are made of the results. These conflicting forces have helped create a burgeoning interest in alternative forms of assessments, particularly complex, performance-based assessments. It is argued that there is a need to rethink the criteria by which the quality of educational assessments are judged, and a set of criteria that are sensitive to some of the expectations for performance-based assessments is proposed.